REVUE DES SCIENCES DE L'EAU, Rev. Sci. Eau 4(1996) 421-434

FREE OPINION

Modeling Hydrologic Phenomena

TRIBUNE LIBRE

Sur la modélisation des phénomènes hydrologiques

J. GANOULIS[1]

Cette tribune libre est publiée intégralement en anglais et en français ; les figures et références bibliographiques, communes aux deux versions, sont placées après les textes.*
This "Free Opinion" is published integrally in both English and French; the figures and references for the two versions can be found at the end of the french text.*

SUMMARY

With the aim of suggesting some practical rules for the use of hydrological models, G. DE MARSILY in his "free opinion" (Rev. Sci. Eau 1994, 7(3): 219-234) proposes a classification of hydrologic models into two categories: (i) models built on data (observable phenomena) and (ii) models without any available observations (unobservable phenomena). He claims that for the former group of observable phenomena, models developed through a "learning process" as well as those based on the underlying physical laws are of the "black box" type. For the latter group of unobservable phenomena, he suggests that physically-based hydrologic models be developed.

Physically-based hydrologic models should introduce to the phenomenological laws the correct empirical coefficients, which correspond to the proper time and space scales (GANOULIS, 1986). Well-known examples are Darcy's permeability coefficient on the macroscopic scale as derived from the Navier-Stokes equations on the local scale and the macroscopic dispersion coefficients in comparison with the local Fickian diffusion coefficients. Misuse of these models by confusing the proper time and space scales and determining the coefficients by calibration is not a sufficient reason to consider them as belonging to the "black box" type. "Black box" type hydrologic models, although very useful when data are available, remain formally empirical. They fail to give correct answers when serious constraints of unity in place, time and action are not fulfilled.

Concerning the second class of models, we may notice that purely "unobservable" phenomena without any available data do not really exist in hydrology. In the case of very rare events and complex systems, such as radioactivity impacts and forecasting of changes on a large scale, physically-based models with adequate para-

1. Department of Civil Engineering, Aristotle University of Thessaloniki, 54006 Thessaloniki, Greece.
Département de Génie Civil, Université Aristote de Thessaloniki, 54006 Thessaloniki, Grèce.

* Les commentaires seront reçus jusqu'au 29 août 1997.

meters may be used to integrate scarce information from experiments and expert opinions in a Bayesian probabilistic framework (APOSTOLAKIS, 1990).

The most important feature of hydrologic models capable of describing real hydrologic phenomena is the possibility of handling imprecision and natural variabilities. Uncertainties may be seen in two categories: aleatory or noncognitive, and epistemic or cognitive. Probabilistic hydrologic models are more suitable for dealing with aleatory uncertainties. Fuzzy logic-based models may quantify epistemic uncertainties (GANOULIS *et al.*, 1996). The stochastic and fuzzy modeling approaches are briefly explained in this "free opinion" as compared to the deterministic physically-based hydrologic modeling.

Key-words : *hydrologic models, black box, physically-based, uncertainties, deterministic, stochastic, fuzzy logic.*

INTRODUCTION

Although this "free opinion" reflects some of the author's general views on the development and use of models in hydrology, it was inspired by the discussion article, "On the Use of Models in Hydrology" by G. DE MARSILY, *Rev. Sci. Eau 1994*, **7**(3): 219-234. Some of the examples which are used for illustration purposes come from the author's experience in water quality and ecological modeling.

From an engineering standpoint, what is important in hydrology is the capability of a hydrologic model to deal with real situations over a relatively large physical space (catchment area) (BOWLES et O'CONNELL, 1991). Hydrologic phenomena on such a scale are complex and dominated by several types of uncertainties (GANOULIS, 1995). For example, the modeling of the rainfall-runoff relationship or the simulation of propagation of flood waves in irregular topography taking into account different hydrologic components, such as evapo-transpiration, infiltration and runoff, show a high degree of variability in space and time. Hydrologic output should contain different uncertainties, making the model predictions imprecise and uncertain. Without losing the general aspects of natural hydrologic phenomena, we can state that a hydrologic model generating a precise output may be suspected to be unrealistic. At the same time, if a model is to describe real hydrologic phenomena, it should be capable of handling imprecision and uncertainty.

G. DE MARSILY, with the aim of suggesting some practical rules, in his aforementioned review paper, proposes a classification of hydrologic models into two groups:

1) *models of observable phenomena*, built on available data, and

2) *models of unobservable phenomena*, which may be developed without any available information.

Different procedures for calibration, validation and use of models belonging to the above categories are suggested. Emphasis is given to the "black box" type models, in which all kinds of empirical or physically-based hydrologic models, such as deterministic, stochastic, conceptual, or distributed, are included.

What is mainly being argued in the present "free opinion" is:

– first, that in the classification of hydrologic models, the distinction between "observable" and "unobservable" phenomena is not the most important feature;

– second, that the combining of physically-based models with empirical constructions of the "black box" type may be misleading;

– third, that the importance of hydrologic modeling to engineering applications is to handle natural and human imprecision/uncertainties, and;

– fourth, that non-utilitarian models may be as important as fundamental research is for engineering applications in understanding and explaining real complex hydrologic phenomena.

OBSERVABLE AND UNOBSERVABLE PHENOMENA

To avoid any confusion, it would be useful to distinguish between various possible kinds of "unobservable" phenomena or entities as follows:

a) Taken literally, "unobservable" entities are out of the scope of natural sciences. In fact, ARISTOTLE distinguished, four centuries before Christ, between those who tried to explain the world by considering "unobservable" or "imaginary" entities (the so-called *theologi)* and those who considered only natural and observable phenomena (the so-called *physici).*

b) An "unobservable" entity should not only be conceived of as one which is not detectable by one of the human senses (such as visual detection), but also by any scientific instrument *currently* available. However, what is today an "unobservable" phenomenon, due to technological limitations, could turn into an observable one in the future, given the development of more sensitive and more accurate scientific instruments.

c) An "unobservable" entity may be considered as one theoretical scientific entity, which may not be measured directly. For example, in groundwater hydrology the notions of pore size, the tortuosity factor and the pore velocity can not be observed or measured as they are. However, they could be estimated indirectly by measuring other entities such as the porosity or the pressure gradient.

None of the above definitions of "unobservable" quantities are part of DE MARSILY'S distinction between "observable" and "unobservable" phenomena. He uses instead the term "unobservable" to denote predictive quantities, both in time and space, for which no data are available. Some examples are the prediction of future radioactivity from nuclear waste disposal in a deep repository or the ecosystem change downstream from Paris due to wastewater treatment upstream.

In fact, what DE MARSILY qualifies as "unobservable phenomena" are situations where direct observations are not available because we are dealing with very *rare events* (low radioactivity in the coming centuries) or very *complex systems* (the effect of phosphate removal upstream of Paris on ecosystem change downstream). This is the usual situation in the safety assessment of complex technological systems (APOSTOLAKIS, 1990) or in risk analysis of environmental and ecological situations (GANOULIS, 1994). What is clear is that in these cases the "black box" type models are not useful because no calibration is possible in the absence of data. However, the use of physically-based models with adequate parameters and the integration of statistical observations and expert opinions in a Bayesian probabilistic framework is the most convenient way to treat the so-called "unobservable phenomena" as defined by DE MARSILY.

This leads us to clearly distinguish between:

a) *"black box"* as opposed to *physically-based* models and

b) *deterministic* as opposed to *stochastic* and *fuzzy logic-based models.*

"Black box" type models are developed using the systems approach (G. DE MARSILY, 1978). The complex natural hydrological system is empirically identified by analysing the relationship between a given set of inputs and the number of outputs it produces. This type of modeling is limited in its use to what DE MARSILY calls the three rules of ancient tragedy, i.e. unity of space, time and action. Although very useful when data is available, they are of very little value in cases where some of the above rules, which are very stringent indeed, are violated or when information is not available. A completely different approach leads us to the physically-based hydrologic models and in fact it is confusing that DE MARSILY still considers them to be the "black box" type. Physically-based hydrological models are developed using a combination between physical principles, such as mass, momentum and energy balance, and empirical laws, such as Darcy's or Fick's Laws (e.g., the "Système Hydrologique Européen" (SHE) (BEVEN *et al.*, 1980) and the USDA Agricultural Research Service Small Watershed Model SWAM (ALONSO et DE COURSEY, 1985), the Canadian Model CEQUEAU (MORIN *et al.*, 1981)). By the use of empirical data, observation and experimentation, these models may be generalized by *induction* in order to adequately describe the complex hydrological processes on the catchment scale. This "inductive" approach based on generalization of observations with physical laws goes back to the foundations of science, when Aristotle formulated his theory of scientific methodology. This "Aristotle approach" for constructing models has been made more dynamic by what French scientist Claude BERNARD defined in 1865 as the three main steps for developing scientific theories (BERNARD, 1966):

1) observation and experimentation,

2) formulation of scientific hypotheses and

3) testing by confirmation or refutation of the hypotheses against the available data.

The advantage of physically-based models over the "black box" type is that they can be used without the severe working conditions of unity in place, time and action. The coefficients of the relevant phenomenological laws should be defined on the proper scale in space and time (GANOULIS, 1986). In this case, very limited calibration relative to the phenomenological coefficients is needed, as for example in simulating coastal circulation and dispersion (GANOULIS, 1991; 1992) or global atmospheric circulation (MATYASZOVSKY *et al.*, 1995). There are of course many difficulties in the proper definition of the empirical coefficients on a large scale, in the appropriate number of these coefficients, and how to take into account several uncertainties (GANOULIS, 1994). However, this is not a reason to consider these models to be of the "black box" type.

We should also stress the importance in hydrology of another class of models, called *analogical models.* Without having direct utility in simulating real-type hydrologic phenomena, these models are based on a simplification of real geometry and change of scale. They are very important in providing insight and understanding into complex hydrologic phenomena, as for example the modeling of a porous medium with a bundle of capillary tubes or random capillary networks (GANOULIS et THIRRIOT, 1977).

MODELING HYDROLOGIC UNCERTAINTIES

Actually, uncertainties are due to our own lack of knowledge about the structure of various physical and biochemical processes and also to the limited amount of data available (DUCKSTEIN et PLATE (eds.), 1987; GANOULIS (ed.), 1991). Several authors in the literature have analyzed different types of uncertainties and made various distinctions, such as between cognitive and noncognitive, objective and subjective, basic and secondary, or natural and technological uncertainties (GANOULIS, 1994). Distinction should be made between:

1) *aleatory* or noncognitive or natural uncertainties or randomness and,
2) *epistemic* or cognitive or man-induced or technological uncertainties.

Aleatory Uncertainties or Randomness

It is postulated that natural uncertainties are inherent to the specific process, mainly because of the space and time variability, and they can not be reduced by use of an improved method or more sophisticated model.

Epistemic or Man-induced Uncertainties

Man-induced uncertainties are of different kinds: (a) data uncertainties, due to sampling methods (statistical characteristics), measurement errors and methods of analyzing the data (b) modeling uncertainties, due to the inadequate mathematical models in use and to errors in parameter estimation, and (c) operational uncertainties, which are related generally to the construction, maintenance and operation of engineering works. Contrary to natural randomness, man-induced uncertainties may be reduced by collecting more information or by improving the mathematical model.

Hydrologic complexity and uncertainties may be quantified by use of:

1) the *stochastic approach* and
2) the *fuzzy set theory*.

Stochastic modeling introduces probabilities in order to simulate hydrological systems, which are subject to uncertainties. In a Bayesian framework, prior information on aleatory uncertainties may be increased into posterior information, by use of additional data (APOSTOLAKIS, 1990). Alternatively, when data are scarce, the fuzzy set theory may be used to handle and quantify imprecision of model parameters and input. The connection between stochastic and deterministic approaches has been analyzed by GANOULIS and MOREL-SEYTOUX (1985) in the case of aquifer systems.

To better understand the methodologies, the stochastic and fuzzy set approaches may be compared with the classical deterministic engineering modeling.

– *Deterministic modeling* provides a *model conditional deterministic solution*. The actual or unconditional or true solution should take into account both aleatory and epistemic uncertainties. The latter are mainly due to model imperfection and parameter uncertainties. In the case of an optimum deterministic model, epistemic uncertainties can be minimized but it is impossible to estimate formally the aleatory uncertainties.

– *Stochastic modeling* allows us to determine the probability distribution law of the solution and furnishes a formal procedure for computing aleatory uncertain-

ties. This can be interpreted as an advantage of the stochastic approach over the deterministic analysis.

– *Fuzzy modeling* is based on fuzzy logic calculus to quantify uncertainties due to imprecision in the values of physical parameters and input boundary conditions (DUBOIS et PRADE, 1980; ZIMMERMANN, 1985; KAUFMANN et GUPTA, 1985). Uncertainties are propagated by combining mathematical modeling with interval operations of different confidence values (DOU *et al.*, 1996; GANOULIS *et al.*, 1996). When data are scarce, output variables such as pollutant concentrations, dissolved oxygen and biomass are calculated directly as fuzzy numbers, without repeating a large number of computations. Fuzzy regression may also be used as an alternative to statistical correlation (GANOULIS et DUCKSTEIN, 1995).

As shown qualitatively in Figure 1, stochastic modeling can minimize all deviations between simulated and real uncertainties as long as the model has sufficient parameters and is improved (in terms of complexity and structure). A deterministic model with sufficient or adequate parameters can minimize deviations on parameter and model uncertainties, but it is not able to handle aleatory uncertainties. Fuzzy logic-based models have similar tendencies and may be considered to be an approximation of stochastic models. Figure 2 provides an example of a fuzzy solution output due to the imprecision on the dispersion coefficient (GANOULIS *et al.*, 1996). All kinds of models fail if the parameters they use are inadequate. In such cases, there may be an optimum for a certain degree of model complexity *(Fig. 1)*. Beyond this optimum, uncertainties increase together with the model's complexity.

CONCLUSION

"Black box" type hydrologic models are essentially empirical and fail to give correct answers when serious constraints of unity in place, time and action are not fulfilled. Very rare events and complex hydrologic systems can be simulated by using physically-based models with adequate coefficients. Available information along with expert opinions can be taken into account in a Bayesian probabilistic framework.

Physically-based hydrologic models should contain empirical coefficients of the phenomenological laws properly defined on large space and time scales. The challenging question is how to evaluate these coefficients by aggregation of the empirical laws, which are valid on a local scale. Estimation of these coefficients only by calibration may result in inaccurate conclusions, mainly as a result of various kinds of hydrologic uncertainties.

Hydrologic uncertainties may be classified into two groups: *aleatory uncertainties* which can be quantified by *stochastic modeling* and *epistemic uncertainties*, which may be estimated with *fuzzy logic*-based modeling. Advantages and disadvantages of different types of hydrologic modeling techniques are analyzed in this "free opinion".

See figures and references after the french text.

RÉSUMÉ

Afin de suggérer quelques règles pratiques sur l'utilisation des modèles hydrologiques, G. De MARSILY dans sa « tribune libre » (Rev. Sci. Eau 1994, 7 (3) : 219-234) propose une classification des modèles hydrologiques en deux catégories : (i) modèles construits sur des données (phénomènes observables) et (ii) modèles sans observations disponibles (phénomènes non-observables). Il préconise que pour la première catégorie des phénomènes observables, les modèles développés selon une « procédure d'apprentissage » ainsi que ceux basés sur les lois physiques sont du type *« boîte noire »*. Pour le deuxième groupe de modèles, il propose l'utilisation de ceux à base physique.

Les modèles hydrologiques *à base physique* doivent introduire dans les lois phénoménologiques des coefficients empiriques corrects vis-à-vis des échelles correspondantes d'espace et de temps (GANOULIS, 1986). Des exemples bien connus sont le coefficient de la perméabilité de Darcy à l'échelle macroscopique, dérivé à partir des équations Navier-Stokes au niveau microscopique et les coefficients de dispersion macroscopique provenant des coefficients de diffusion locale de Fick.

Une utilisation abusive de ces modèles en introduisant des échelles de temps et d'espace inappropriées et en déterminant les coefficients par étalonnage, n'est pas une raison suffisante pour les considérer de type « boîte noire ». Les modèles hydrologiques du type « boîte noire » sont très utiles lorsque des données sont disponibles, mais restent essentiellement empiriques. Ils ne donnent pas de réponses correctes lorsque des restrictions sévères d'unité de lieu, de temps et d'action ne sont pas valables.

Concernant la deuxième classe des modèles, nous pouvons remarquer que dans la réalité des phénomènes purement « non observables » et sans aucune donnée, n'existent pas en hydrologie. Dans les cas *d'événements rares* et de systèmes hydrologiques complexes tels que les impacts dus à la radioactivité et les prévisions à grande échelle, les modèles á base physique avec des paramètres adéquats peuvent être utilisés afin d'intégrer les informations rares et les opinions des experts, dans un cadre probabiliste Bayesien (APOSTOLAKIS, 1990).

La propriété la plus intéressante des modèles hydrologiques est leur capacité éventuelle à tenir compte des imprécisions et des variabilités naturelles. Les incertitudes peuvent être classées en deux catégories : *aléatoires* (liées à la variabilité naturelle) et *épistémiques* (liées à l'approche scientifique). Des modèles hydrologiques probabilistes sont plus convenables pour tenir compte des incertitudes aléatoires. Des modèles basés sur la logique floue peuvent quantifier les incertitudes épistémiques (GANOULIS *et al.*, 1996). Les approches stochastiques et flous sont brièvement expliquées dans cette « tribune libre » en comparaison avec les modèles hydrologiques déterministes à base physique.

Mots clés : *modèles hydrologiques, boîte noire, incertitudes, stochastique, logique floue.*

INTRODUCTION

Quoique cette « tribune libre » reflète l'opinion de l'auteur sur le développement et l'utilisation des modèles en hydrologie, elle est largement inspirée de l'article de G. DE MARSILY « Sur l'utilisation des modèles hydrologiques », Rev. Sci. Eau 1994, **7** (3) : 219-234. Quelques uns des exemples utilisés afin d'illustrer des arguments proviennent de l'expérience de l'auteur sur la modélisation écologique et de qualité des eaux.

Du point de vue de l'ingénieur, ce qui est important en hydrologie est la capacité des modèles hydrologiques de représenter des situations réelles dans un espace physique relativement large (échelle du bassin versant). A cette échelle, les phénomènes hydrologiques sont complexes et dominés par plusieurs types d'incertitudes. Par exemple, la modélisation de la relation pluie-débit ou la simulation de la propagation des ondes de crue sur une topographie irrégulière compte tenu des diverses composantes hydrologiques, comme l'évapotranspiration, l'infiltration et l'écoulement superficiel, possèdent un grand degré de variabilité dans l'espace et le temps. Les résultats des modèles hydrologiques contiennent plusieurs incertitudes qui font toute prédiction imprécise et incertaine. Sans perdre de vue les aspects généraux des phénomènes hydrologiques physiques, on peut dire que plus les modèles hydrologiques produisent des résultats précis, plus ils sont suspects d'être loin de la réalité. En même temps, si un modèle hydrologique est fait pour décrire des phénomènes hydrologiques réels, il doit être capable de tenir compte de l'imprécision et de l'incertitude.

Suggérant quelques règles pratiques, G. DE MARSILY, à la tribune libre citée ci-dessus, propose une classification des modèles hydrologiques en deux groupes :

1) *modèles des phénomènes observables*, qui sont mis au point à partir de données ;

2) *modèles des phénomènes non observables*, qui sont développés sans aucune information préalable.

Diverses procédures pour l'étalonnage, la validation et l'utilisation des modèles appartenant aux catégories ci-dessus sont suggérées. L'accent est mis sur les modèles du type « boîte noire » dans lesquels sont inclus tout genre de modèles : empiriques, à base physique, déterministes, stochastiques, conceptuels ou distribués.

Dans cette « tribune libre », les principaux arguments présentés sont :

1) dans la classification des modèles hydrologiques, la distinction entre « observable » et « non observable » n'est pas la notion la plus importante ;

2) combiner les modèles à base physique avec les modèles empiriques du type « boîte noire » peut conduire à une confusion ;

3) l'importance de la modélisation hydrologique aux applications de l'ingénieur consiste en la possibilité de maîtriser les imprécisions/incertitudes naturelles ou humaines ;

4) que certains modèles non-utilitaires peuvent être aussi importants, comme la recherche fondamentale est aux applications de l'ingénieur, afin de mieux comprendre et expliquer des phénomènes hydrologiques complexes.

PHÉNOMÈNES OBSERVABLES ET NON OBSERVABLES

Afin d'éviter toute confusion, il serait utile de distinguer entre plusieurs types de phénomènes ou entités « non-observables ».

a) Selon son sens premier, une entité « non-observable » ne relève pas des sciences physiques. En effet, quatre siècles avant J.-C., ARISTOTE faisait la distinction entre ceux qui essayaient d'expliquer le monde en considérant des enti-

tés non-observables ou imaginaires (nommés *theologi)* et ceux qui considéraient seulement des phénomènes physiques et observables (nommés *physici).*

b) Comme entité « non-observable » doit être considérée, non seulement une entité qui ne puisse pas être détectée par les sens humains (comme la vision) mais par un quelconque instrument scientifique, actuellement disponible. Toutefois, ce qui aujourd'hui est un phénomène « non-observable » à cause des limites technologiques, peut devenir « observable » dans l'avenir, étant donné la mise au point d'instruments scientifiques plus sensibles et plus perfectionnés.

c) Une entité « non-observable » peut être considérée comme une entité scientifique théorique, qui ne soit pas directement mesurable. Par exemple, en hydrologie souterraine, les notions physiques de taille d'un pore, du facteur de tortuosité ou de la vitesse au niveau du pore ne peuvent pas être mesurées ou observées en tant que telles. Toutefois, elles peuvent être estimées indirectement en mesurant d'autres grandeurs physiques comme la porosité et la différence de pression.

Aucune des définitions ci-dessus ne correspond à la distinction que fait DE MARSILY entre phénomènes « observables » et « non observables ». Il utilise plutôt le terme « non observable » pour décrire des quantités estimées dans le temps et l'espace, pour lesquelles il n'y a pas de données disponibles. De tels exemples sont la prévision de la radioactivité émise par des déchets nucléaires ou l'impact sur les écosystèmes à l'aval de la ville de Paris dû à la déphosphatation des eaux-usées en amont.

En réalité, ce que DE MARSILY qualifie de phénomènes non observables sont des situations ou les observations directes ne sont pas disponibles parce qu'on a à faire à des événements *très rares* (faible radioactivité dans un avenir lointain) ou à des *systèmes très complexes* (écosystème en aval de la ville de Paris). C'est le cas de l'analyse de la sûreté des systèmes technologiques complexes (APOSTOLAKIS, 1990) ou de l'analyse du risque environnemental et écologique (GANOULIS, 1994). Il est évident que dans ces cas-là les modèles du type « boîte noire » ne sont pas utiles, car, faute de données, ils ne peuvent pas être étalonnés. Par contre, l'utilisation des modèles basés sur les lois physiques avec les paramètres adéquats et la juxtaposition des observations statistiques et des opinions expertes dans un cadre probabiliste Bayesien, est la façon la plus convenable de traiter les phénomènes que DE MARSILY qualifie comme non observables.

Ceci nous amène à distinguer clairement entre :

a) les modèles du type *« boîte noire »* par rapport à ceux *à base physique* et

b) les modèles *déterministes* par rapport aux modèles *stochastiques* et ceux *à base de logique floue.*

Les modèles « boîte noire » sont développés selon l'approche systémique (G. DE MARSILY, 1978). La structure complexe du système hydrologique est identifiée empiriquement en analysant la relation entre un ensemble donné des entrées-sorties. Ce genre de modèles est limité dans son utilisation par ce que DE MARSILY qualifie les trois règles de la tragédie antique, à savoir l'unité de lieu, de temps et d'action. Quoique ces modèles soient très utiles lorsque des données sont disponibles, ils sont d'une valeur limitée en l'absence d'informations préalables. Une approche complètement différente conduit aux modèles hydrologiques basés sur les lois physiques et le fait que DE MARSILY continue à appeler ces modèles « boîtes noires » prête à confusion.

Les modèles hydrologiques à base physique sont développés en combinant des lois physiques telles que la conservation de la masse, de la quantité du mouvement et de l'énergie aux lois phénoménologiques, comme les lois de Darcy et de Fick (voir le Système Hydrologique Européen (SHE) – BEVEN *et al.*, 1980, le modèle SWAM du Service de la Recherche de l'Agriculture Américaine – (ALONSO et DE COURSEY, 1985) et le modèle canadien CEQUEAU (MORIN *et al.*, 1981). En utilisant des données empiriques, l'observation et l'expérimentation, ces modèles peuvent être généralisés par « induction », afin de décrire correctement les processus hydrologiques complexes à l'échelle du bassin. Cette approche « inductive » basée sur les lois physiques et la généralisation des observations, nous ramène aux frontières de la science, lorsque ARISTOTE formulait sa théorie de la méthodologie scientifique. Cette approche d'ARISTOTE a trouvé un nouveau dynamisme par ce que le scientifique français C. BERNARD définissait en 1865 comme les trois étapes pour développer des théories scientifiques (BERNARD, 1966) :

1) observation et expérimentation ;
2) formulation des hypothèses scientifiques ;
3) contrôler par confirmation ou rejet la validation des hypothèses en tenant compte desdonnées disponibles.

L'avantage des modèles à base physique par rapport à ceux du type « boîte noire » réside dans le fait qu'ils peuvent être utilisés sans les contraintes sévères de l'unité de lieu, de temps et d'action. Les coefficients qui interviennent dans les lois phénoménologiques doivent être définis à la juste échelle de l'espace et du temps (GANOULIS, 1986). Dans ce cas, un étalonnage très limité est nécessaire, comme par exemple pour la simulation de la circulation et de la dispersion côtière (GANOULIS, 1991, 1992) ou de la circulation atmosphérique globale (MATYASZOVSKY *et al.*, 1995). Bien sûr, il y a plusieurs difficultés sur la définition des coefficients empiriques à grande échelle, le grand nombre de ces coefficients à prendre en compte et des diverses incertitudes (GANOULIS, 1994). Toutefois, il n'y a pas de raison pour considérer ces modèles comme étant du type « boîte noire ».

Nous allons souligner aussi l'importance en hydrologie d'une autre classe de modèles, connus comme *modèles analogiques*. Sans avoir une utilité directe pour simuler des phénomènes hydrologiques, ces modèles sont basés sur une simplification de la géométrie réelle et un changement d'échelle. Ils sont très importants pour la compréhension et l'approfondissement des phénomènes hydrologiques complexes, comme par exemple la modélisation du milieu poreux par un faisceau de tubes capillaires ou par des réseaux de capillaires interconnectés au hasard (GANOULIS et THIRRIOT, 1977).

MODÉLISATION DES INCERTITUDES HYDROLOGIQUES

Les incertitudes sont dues surtout à notre ignorance de la structure des divers processus physiques et biochimiques et aussi à la quantité limitée des données (DUCKSTEIN et PLATE (eds.) 1987, GANOULIS (ed.) 1991). Dans la littérature, plusieurs auteurs ont analysé les différents types d'incertitude, faisant distinction entre incertitudes cognitives et non cognitives, objectives et subjectives, de base et secondaires, naturelles ou technologiques (GANOULIS, 1994).

Nous pouvons distinguer entre incertitudes :
1) *aléatoires* ou *non cognitives* ou *naturelles* et
2) *épistémiques* ou *induites par l'homme* ou *technologiques.*

Incertitudes aléatoires : il est postulé que les incertitudes naturelles sont liées au processus physiques surtout à cause de la variabilité dans l'espace et le temps. Elles ne peuvent pas être réduites en utilisant une méthode améliorée ou un modèle plus sophistiqué.

Incertitudes épistémiques : les incertitudes induites par l'homme sont diverses (a) incertitudes sur les données, celles dues aux méthodes d'échantillonnage (caractéristiques statistiques), erreurs de mesure et méthodes d'analyse des données (b) incertitudes de modélisation dues aux modèles mathématiques inadéquats et aux erreurs d'estimation des paramètres et (c) incertitudes opérationnelles qui sont en général liées à la construction, maintenance et mode d'opération des travaux de l'ingénieur. Contrairement aux incertitudes aléatoires, celles induites par l'homme peuvent être réduites en recueillant plus d'informations et en améliorant le modèle mathématique.

Les incertitudes hydrologiques peuvent être quantifiés en utilisant
1) l'approche *stochastique* et
2) la théorie des *ensembles flous.*

La modélisation stochastique introduit des probabilités afin de simuler les systèmes hydrologiques soumis aux incertitudes. Dans un cadre Bayesien, l'information préalable sur les incertitudes aléatoires peut être augmentée au niveau de l'information postérieure en utilisant des données supplémentaires (APOSTOLAKIS, 1990). Alternativement, quand les données sont rares, la théorie des ensembles flous peut être utilisée afin de contrôler et de quantifier l'imprécision sur les paramètres du modèle et les conditions limites. La relation entre les approches stochastiques et déterministes a été analysée par GANOULIS et MOREL-SEYTOUX (1985) dans le cas des systèmes aquifères.

Pour mieux comprendre les méthodologies, les approches stochastiques et celles basées sur des ensembles flous peuvent être comparées à la modélisation déterministe classique de l'ingénieur.

– *La modélisation déterministe* conduit à la *solution déterministe conditionnée au modèle.* La solution exacte ou non-conditionnée doit tenir compte aussi bien des incertitudes aléatoires qu'épistémiques. Ces dernières sont dues à l'imperfection du modèle et aux incertitudes sur les paramètres. Dans le cas d'un modèle déterministe optimal, les incertitudes épistémiques peuvent être minimisées mais il est impossible d'estimer les incertitudes aléatoires.

– *La modélisation stochastique* permet de déterminer la loi de distribution probabiliste de la solution et prévoit une procédure formelle pour calculer les incertitudes aléatoires. Ceci peut être interprété comme un avantage de l'approche stochastique par rapport à l'analyse déterministe.

– *La modélisation floue* est basée sur l'arithmétique floue (DUBOIS et PRADE, 1980 ; ZIMMERMANN, 1985 ; KAUFMANN et GUPTA, 1985), afin de quantifier les incertitudes dues à l'imprécision sur la valeur des paramètres physiques et les conditions aux frontières. Les incertitudes sont propagées en combinant la modélisation mathématique déterministe aux opérations sur les intervalles à valeur de confiance variable (DOU *et al.*, 1996 ; GANOULIS *et al.*, 1996). Lorsque les données sont rares, les variables de sortie, telles que la concentration du polluant, l'oxygène dissoute et la biomasse sont calculées en tant que nombres flous, sans

répéter un grand nombre de calculs. La régression floue peut être aussi utilisée comme alternative à la corrélation statistique (GANOULIS et DUCKSTEIN, 1995).

Comme le montre la figure 1, la modélisation stochastique peut minimiser tous les écarts entre incertitudes simulées et réelles, à condition que le modèle ait suffisamment de paramètres et soit plus amélioré (au sens de sa structure). Un modèle déterministe avec des paramètres suffisants et adéquats peut minimiser les écarts d'incertitudes dus aux paramètres et au modèle, mais il n'est pas capable de tenir compte des incertitudes aléatoires. Les modèles basés sur la logique floue ont des tendances similaires et peuvent être considérés comme une approximation des modèles stochastiques. La figure 2 donne un exemple de résultat de calcul flou dû à l'imprécision du coefficient de la dispersion (GANOULIS *et al.*, 1996). Tous modèles ne sont pas corrects lorsque leurs paramètres sont inadéquats. Dans ce cas, il y a un optimum pour un certain degré de complexité du modèle. En dehors de cet optimum, la précision du modèle croît en même temps que sa complexité.

CONCLUSION

Les modèles hydrologiques du type « *boîte noire* » sont essentiellement empiriques et ne peuvent donner de réponses correctes en absence de restrictions rigoureuses d'unité de lieu, de temps et d'action.

Des événements très complexes et très rares peuvent être simulés à l'aide des modèles *à base physique* et ayant des coefficients adéquats. L'information disponible aussi bien que les opinions des experts peuvent être pris en compte dans un cadre probabiliste Bayesien.

Les modèles hydrologiques à base physique doivent introduire les coefficients des lois phénoménologiques définis d'une manière appropriée à grande échelle d'espace et de temps. La question difficile est comment évaluer ces coefficients par agrégation des lois empiriques valables à l'échelle locale. Lorsque l'estimation de ces coefficients se fait seulement par étalonnage, on obtient des conclusions approximatives, surtout à cause des incertitudes hydrologiques. Ces dernières peuvent être classifiées en deux groupes :

(1) incertitudes *aléatoires*, qui peuvent être mieux quantifiées par modélisation stochastique et (2) incertitudes *épistémiques* qui peuvent être estimées par modélisation basée sur la logique floue. Les avantages et inconvénients des différents types de modélisation sont analysées dans cette tribune libre.

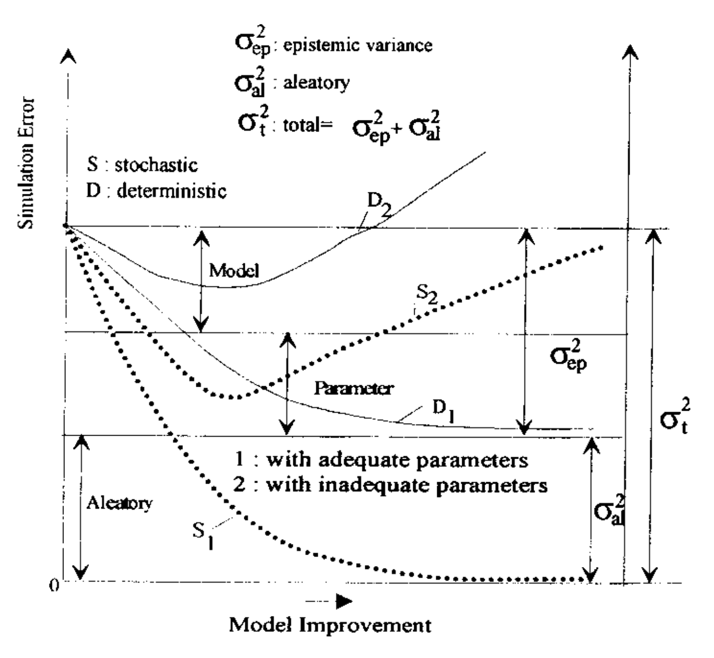


Figure 1 Relation qualitative entre erreurs de simulation et amélioration de la modélisation pour différents types de modèles.

Qualitative relation between simulation errors and model improvement for different types of models.

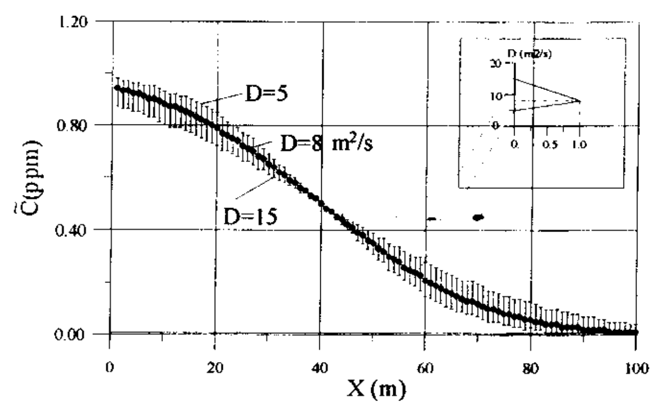


Figure 2 Distribution du $\tilde{C}$ flou en dispersion convective unidimensionnelle.

Distribution of fuzzy $\tilde{C}$ for one-dimensional convective dispersion.

RÉFÉRENCES BIBLIOGRAPHIQUES

ALONSO C.V., DE COURSEY D.G., 1985. Small Watershed Model, D.G. de Coursey (ed.), *Proc. of the Natural Resources Modeling Symposium*, Pingree Park, CO, USDA-ARS-30, 40-46.

APOSTOLAKIS G., 1990. The concept of probability in safety assessments of technological systems. *Science* **250** : 1359-1364.

BERNARD C., 1966. *Introduction à l'étude de la médecine expérimentale*, Garnier-Flammarion, Paris, 318 pp.

BEVEN K.J., WARREN R., ZAOUI J., 1980. SHE : Towards a Methodology for Physically-Based, Distributed Modeling in Hydrology, *Hydrological Forecasting*, Int'l Ass. Hydrol. Sci., Pub. N° 129, 133-137.

BOWLES D., O'CONNELL P.E. (eds.), 1991. *Recent Advances in the Modeling of Hydrologic Systems*, NATO ASI Series C, Vol. 345, Kluwer Academic, Dordrecht, The Netherlands, 667 pp.

DOU C., WOLDT W., BOGARDI I., DAHAB M., 1996. Steady-State Groundwater Flow Simulation with Imprecise Parameter, *Wat. Resour. Res.* (to appear).

DUBOIS D., PRADE H., 1980. *Fuzzy Sets and Systems : Theory and Applications*, Academic Press, New York.

DUCKSTEIN L., PLATE E. (eds.), 1987. *Engineering Reliability and Risk in Water Resources*, E.M. Nijhoff, Dordrecht, The Netherlands, 588 pp.

GANOULIS J., 1986. Sur les échelles spatiales des hétérogénéités en milieu poreux, *Hydrogéologie*, N° 2 : 115-123, BRGM, France.

GANOULIS J. (ed.), 1991. *Water Resources Engineering Risk Assessment*, NATO ASI Series, Vol. G 29, Springer-Verlag, Heidelberg, Germany, 539 pp.

GANOULIS J., 1991. Water Quality Assessment and Protection Measures of a Semi-enclosed Coastal Area : the Bay of Thermaikos, *Marine Pollution Bulletin*, **23**, 83-87.

GANOULIS J., 1992. Dispersion and disappearance of coliform bacteria in the bay of Thessaloniki, *Rev. Sci. Eau*, **5** (4): 541-554.

GANOULIS J., 1994. *Engineering Risk Analysis of Water Pollution : Probabilities and Fuzzy Sets*, VCH, Weinheim, Germany, 306 pp.

GANOULIS J., 1995. Floodplain Protection and Management in Karst Areas, In : *Defence from Floods and Floodplain Management*, Gardiner J. *et al.* eds., NATO ASI Series, Vol. 299 : 419-428, Kluwer Academic, Dordrecht, The Netherlands.

GANOULIS J., THIRRIOT C., 1976. Stochastic study of the entrance of a non wetting fluid in porous medium, under the influence of the capillarity, *Hydraulic Problems Solved by Stochastic Methods*, Hjorth, *et al.* (eds.), 433-450, Water Resources Publications, Colorado, USA

GANOULIS J., MOREL-SEYTOUX H., 1985. Application of stochastic methods to the study of Aquifer Systems, *Technical Documents in Hydrology*, UNESCO, Paris.

GANOULIS J., DUCKSTEIN L., 1995. Fuzzy Regression between Atmospheric and Coastal Water Temperature for Ecological Modeling, *Pure and Applied Mathematics*, **3** : 59-76.

GANOULIS J., BIMBAS I., DUCKSTEIN L., BOGARDI I., 1996. Fuzzy Arithmetic for Ecological Risk Management, *Risk-Based Decision-Making in Water Resources VII*, Y. Haimes, D. Moser, E. Stakhiv (eds.), (in press).

KAUFMANN A., GUPTA M., 1985. *Introduction to Fuzzy Arithmetic : Theory and Applications*, Von Nostrand Reinhold, New York.

MARSILY G. De, 1978. *De l'Identification des Systèmes Hydrogéologiques*, Thèse d'Etat, Paris VI.

MATYASOVSZKY I., BOGARDI I., GANOULIS J., 1995. Impact of Global Climatic Change on Temperature and Precipitation in Greece, *Appl. Math. Comp.*, **70** : 1-30, 1995.

MORIN G., FORTIN J.-P., SOCHANSKA W., PAQUETTE S., 1981. Modèle CEQUEAU : manuel d'utilisation, INRS-Eau, *rapport scientifique* no. 93, 449 pp.

ZIMMERMANN H.J., 1985. *Fuzzy Set Theory and its Application*, Martinus Nijhoff, Dordrecht, The Netherlands.